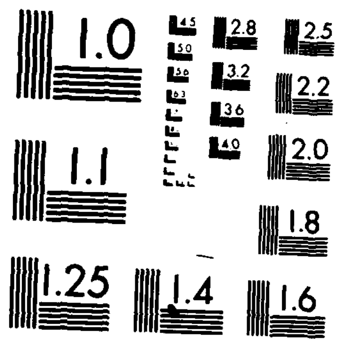
AD-A170 563 THE EFFECTS OF HIGH SUSTAINED ACCELERATION ON THE ACOUSTIC PHONETIC STRUC..(U) HARRY G ARMSTRONG AEROSPACE MEDICAL RESEARCH LAB WRIGHT-PATTE.. 1/1

UNCLASSIFIED Z S BOND ET AL. 21 MAY 86 AAMRL-TR-86-011 F/G 5/7 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

AAMRL-TR-86-011

THE EFFECTS OF HIGH SUSTAINED ACCELERATION ON THE ACOUSTIC PHONETIC STRUCTURE OF SPEECH: A PRELIMINARY INVESTIGATION

ZINNY S. BOND
THOMAS J. MOORE
TIMOTHY R. ANDERSON

MAY 1986

Approved for public release; distribution unlimited.

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

NOTICES

TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-86- 011

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Regulation 169-3.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Henning E. von Gierke

HENNING E. VON GIERKE, Dr Ing
Director
Biodynamics and Bioengineering Division
Air Force Aerospace Medical Research Laboratory

AD-A170 563

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): AAMRL-TR-86-011	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Harry G. Armstrong Aerospace Medical Research Laboratory; 6b. OFFICE SYMBOL (If applicable): AAMRL/BBA	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): Wright-Patterson AFB OH 45433	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2312	V3	37

11. TITLE (Include Security Classification) The Effects of High Sustained Acceleration on the Acoustic Phonetic structure of Speech: A Preliminary Investigation

12. PERSONAL AUTHOR(S) Zinny S. Bond, Thomas J. Moore and Timothy R. Anderson

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	86 May 21	

16. SUPPLEMENTARY NOTATION
AAMRL/Contact: Thomas J. Moore, AAMRL/BBA, Tel: (513) 255-3607.

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Stressed Speech; Acceleration; Acoustic-Phonetics

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The acoustic-phonetic structure of speech produced under adverse circumstances such as high levels of noise, vibration, and stress, has received little investigation. The purpose of this study was to provide some preliminary data concerning speech produced under high sustained acceleration. Acoustical measurements were made of a set of words spoken by two subjects at 1 G and +6 Gz. Words produced under acceleration differed from words produced at 1 G in both spectral and durational characteristics.

The formant shifts observed were similar for both speakers. The first formant increased for the majority of vowels. The second formant tended to be lower for the front vowels /i, ɨ/ and higher for the back vowel (u). Acceleration resulted in a raising of the first formant and a lowering of the second formant relative to the same diphthongs produced at 1 G.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Thomas J. Moore	22b. TELEPHONE NUMBER (Include Area Code): (513) 255-3607; 22c. OFFICE SYMBOL: AAMRL/BBA

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

The effects of acceleration on word duration differed for the two speakers. One produced almost all words under acceleration at a greater mean duration; for the other speaker, some words increased in duration, others did not. Mean vowel duration followed the same pattern, increasing for almost all vowels for one speaker and not for the second speaker. Consonants were shorter in words produced under acceleration for both speakers.

ABSTRACT

The acoustic-phonetic structure of speech produced under adverse circumstances such as high levels of noise, vibration, and stress, has received little investigation. The purpose of this study was to provide some preliminary data concerning speech produced under high sustained acceleration. Acoustical measurements were made of a set of words spoken by two subjects at 1 G and +6 Gz. Words produced under acceleration differed from words produced at 1 G in both spectral and durational characteristics.

The formant shifts observed were similar for both speakers. The first formant increased for the majority of vowels. The second formant tended to be lower for the front vowels /i, ɛ / and higher for the back vowel (u). Acceleration resulted in a raising of the first formant and a lowering of the second formant relative to the same diphthongs produced at 1 G.

The effects of acceleration on word duration differed for the two speakers. One produced almost all words under acceleration at a greater mean duration; for the other speaker, some words increased in duration, others did not. Mean vowel duration followed the same pattern, increasing for almost all vowels for one speaker and not for the second speaker. Consonants were shorter in words produced under acceleration for both speakers.

PREFACE

This research was accomplished in the Biological Acoustics Branch, Biodynamics and Bioengineering Division, Armstrong Aerospace Medical Research Laboratory, Aerospace Medical Division (AMD). The effort was accomplished under Work Unit 2312V337, "Auditory Information Processing."

The research was sponsored by the Air Force Office of Scientific Research/AFSC, United States Air Force, under Contract F49620-85-C-0013.

The first author is a faculty member at Ohio University, Athens, OH and was a Summer Fellow in the 1985 USAF-UES Summer Faculty Research Program.

The authors wish to acknowledge the advice of Dr. Charles W. Nixon during the preparation of this report.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The purpose of this report is to present data concerning the effects of high sustained acceleration on the acoustic-phonetic structure of speech. These data are of some interest in themselves, in that they provide information about human speech production under adverse circumstances and thus aid in understanding human speech production in general. The data also have practical value. In recent years the Air Force and the Navy have expressed increasing concern over the geometric increase in pilot workload encountered in modern high-performance aircraft. Among the technologies proposed to alleviate this workload is that of automatic speech recognition. The concept has been advanced that many tasks, e.g., radio channel selection, selection of displays and status indicators, etc., that presently require the pilot to manipulate switches, knobs and buttons and/or divert his attention from the outside world into the cockpit could be accomplished by voice input. If automatic speech recognition systems are ever to be usefully employed in aircraft, these systems will have to be able to recognize speech produced under many types of adverse circumstances, e.g., noise, stress, vibration, etc. Acceleration is one of the adverse circumstances for speech production that has to be investigated. If systematic changes occur in speech production under acceleration, compensatory algorithms can be developed to increase the probability that automatic speech recognition

technology will be successfully applied in the dynamic flight environment.

BACKGROUND

Accelerations are classified according to the direction in which they act on the human body by a three-axis (x, y, z) coordinate system, as shown in Figure 1 (taken from Davis, et al., 1984). Headwards acceleration tends to displace body tissue footward; the resultant force is termed positive G or +Gz. High sustained acceleration is defined by Burton, Leverett, and Michaelson (1974) as exposure to acceleration forces of 6 G or greater for periods in excess of 15 seconds.

The physiological effects of exposure to high sustained acceleration have been investigated quite extensively (Sharp and Ernsting, 1978, provide a review). Of these effects, two seem to be particularly relevant to speech production: changes in respiration and muscle control. According to Gillies (1965), positive acceleration produces "little respiratory embarrassment, at least at levels compatible with consciousness... The first [symptom] ... is a slight difficulty with inspiration as the thorax has to be lifted against a greater gravitational force, and for the same reason the expiratory phase becomes more rapid" (p. 618). Respiratory symptoms would be expected to be present at +5 or +6 Gz.

Acceleration has effects on mobility involving large muscles at even relatively low levels. It is impossible to rise from the

sitting position at +3 Gz, for example. However, fine motor control movements can be performed with little or no loss of accuracy at accelerations of at least +8 Gz (Sharp and Ernsting, 1978).

In order to maintain vision and consciousness at higher accelerations, some anti-G straining maneuvers are necessary. These involve pulling the head down, tensing the skeletal and abdominal muscles as much as possible, and increasing intrathoracic pressure by forcibly exhaling against a partially or completely closed glottis (Burton, et al., 1974). Since these straining maneuvers undoubtedly have an effect on vocal tract configuration, they may also have some effect on speech quality.

This investigation examined speech samples produced by two speakers in normal circumstances and at high sustained acceleration, to obtain preliminary information about the effects of acceleration on the acoustic-phonetic structure of speech.

SUBJECTS

The speakers were male, volunteer members of the Acceleration Hazardous Duty Panel[1]. The subjects were permitted to participate in the experiment only upon the approval of an attending medical officer.

The speaking style of the two men sounded quite different at first hearing. The first speaker, identified as #1, spoke more

rapidly and at a higher perceived pitch than the second speaker. The second speaker, #2, had a pronounced tendency to laryngealize the final portions of the last syllable of words, producing some of these with extremely low vocal fold vibration rates. For example, he ended one token of the word two at a fundamental frequency of 40 Hz.

GENERATION OF SPEECH MATERIALS

The recordings of speech under acceleration were produced using the centrifuge at the Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio[2]. The recordings were made using a TEAC Tascam 44 4-channel tape recorder and an M-101 noise-cancelling military microphone. The speakers were wearing standard Air Force oxygen masks and breathing air was supplied through a chest mounted regulator. Since the microphone is located in the oxygen mask, the recordings contain the sounds of breathing which sometimes overlap the onset of words and create difficulties in locating their beginnings, particularly of words with initial fricatives. Sometimes the speech signal itself was noisy.

Each speaker recorded five tokens of each item from a list of 15 Air Force vocabulary words while sitting in the gondola of the centrifuge without acceleration. Each speaker then recorded several tokens of each word in random order while at an acceleration level of +6 Gz. The number of tokens recorded at +6 Gz was determined by the physical state of the subject. Subjects

were permitted to remain at +6 Gz acceleration for 30 seconds at a time with intervening rest periods, for a total exposure time at +6 Gz of three minutes per day. If a subject had no significant objective or subjective evidence of fatigue, he was permitted to add 90 seconds to the three minute daily total time. Consequently, the number of tokens of each word recorded at +6 Gz varied from word to word and from subject to subject.

From these recordings, five words were selected for analysis. The words varied in length from one to four syllables and contained a number of different vowels and consonants. However, the recorded vocabulary did not permit a detailed analysis of a large portion of English segments. The words selected for analysis were: two, seven, zero, frequency, and the letter sequence CCIP, pronounced as /sisiaipi/. The number of tokens of each word which could be analyzed for both subjects and conditions is given in Table 1.

METHOD

Each token of each word was digitized at 16 kHz using a 6.4 kHz anti-aliasing filter and 16 bit resolution; each token was stored on disk. Recording and analysis was performed using SPIRE[3], on the Symbolics 3670 computer. Figure 2 is a display created with SPIRE of the word two, as produced by speaker 2 at +6 Gz. The SPIRE display shows a wide-band spectrogram, formants, the wave-form, and the Linear Predictive Coder (LPC) spectrum.

Segment boundaries were located from a wide-band spectrogram display and a wave-form display, using the segmentation criteria suggested by Peterson and Lehiste (1960). At times, segmentation of the test words was difficult or indeterminate because the segment boundaries themselves were indistinct or because of the recording quality and the occasional presence of noise.

Measurements were made of the first three formants of all identifiable vowels. The measurements consisted of the duration of each word, the duration of each vowel with a clearly defined boundary, the duration of all intervocalic obstruents, the fundamental frequency, and amplitude within low and high frequency bands at the mid-point of each syllable. In addition, LPC spectrum slices were obtained of stressed vowels.

Values for the first three formants were obtained at the mid-point of each vowel from the formant display provided by SPIRE. The formant values for the first component of the diphthong /ai/ were measured at the point at which F2 reached its lowest value; the point at which it reached its highest value was assigned to the second component of the diphthong. For the diphthong /oʊ/, formant values were measured at the point at which F3 reached a steady-state value and at the end of the word. The onset of the diphthong /oʊ/ was quite difficult to specify with precision.

Fundamental frequency was obtained at the midpoint of each syllable from the duration of a glottal period displayed on a waveform. Amplitude was measured in a low-frequency band (125 Hz to 440 Hz) and a high-frequency band (3400 Hz to 5000 Hz).

Table 1. Number of tokens of the test words analyzed for each speaker and condition.

	Speaker 1		Speaker 2	
Word	1G	+6Gz	1G	+6Gz
two	5	3	5	2
seven	4	3	5	4
zero	5	3	5	3
frequency	5	3	3	4
CCIP	5	3	5	4

RESULTS

While speaking under acceleration at +6 Gz, both speakers sounded intelligible and natural, though other portions of the recording suggested that they were experiencing considerable physical stress and were practicing anti-G straining maneuvers to maintain vision. It is not clear whether speech produced under acceleration could be identified as such without other information. However, measurable differences between speech produced at 1 G and at +6 Gz could be found for both speakers.

Vowel formants. Acceleration appears to affect the vowel formants in essentially the same way for both speakers even though their vowel spaces are different. The average values for the first three formants of the vowels / i, ɛ , u/ are given in Tables 2 and 3.

The first formant is raised for all three vowels for speaker 1 and for /i/ and /ɛ/ for speaker 2. That of /u/ is lowered for speaker 2. The second formant of /i/ and /ɛ/ is lowered and that

of /u/ is raised for both speakers. The values of the third formant do not appear to be systematically affected by sustained acceleration for either speaker.

These changes in formant patterns of vowels for both speakers suggest that the effects of acceleration are to compact the vowel space by shifting the formants of the vowels towards a centralized position. This is observed in Figures 3a and 3b where each stressed vowel is located in a space defined by F1 and F2. In articulatory terms, this formant shift would imply reduced displacement of the articulators. The effect is most clearly present for the second formant which is traditionally associated with tongue advancement.

Diphthong formants. Acceleration appears to influence /ai/ and /oμ/ diphthong formant patterns of both talkers in a similar way as shown in Tables 4 and 5. The first formant is raised and the second formant is lowered for both components of /ai/ and /oμ/. The value of the third formant of these diphthongs does not appear to be systematically shifted in speech produced under acceleration.

Table 2. Means and standard deviations, in Hz, of the first three formants for vowels while speaking in 1 G and +6 Gz conditions (speaker 1).

1 G

Vowel	Tokens (N)	F1	F2	F3
ε	4	558 (36)	1356 (47)	2232 (183)
u	5	372 (0)	1308 (121)	2145 (259)
i	20	382 (30)	1849 (129)	2375 (222)

+6 Gz

ε	3	589 (31)	1343 (65)	2356 (112)
u	3	413 (18)	1447 (72)	1984 (112)
i	12	403 (30)	1749 (139)	2264 (121)

Table 3. Means and standard deviations, in Hz, of the first three formants for vowels while speaking in 1 G and +6 Gz conditions (speaker 2).

1 G

Vowel	Tokens (N)	F1	F2	F3
ε	5	539	1358	2189
		(28)	(40)	(17)
u	5	391	967	2071
		(17)	(113)	(26)
i	18	367	1913	2323
		(19)	(66)	(215)

+6 Gz

Vowel	Tokens (N)	F1	F2	F3
ε	4	574	1217	2240
		(18)	(96)	(85)
u	2	357	1209	2093
		(22)	(88)	(22)
i	15	387	1747	2238
		(28)	(92)	(257)

Table 4. Means and standard deviations, in Hz, of initial and final portions of diphthongs produced in 1 G and +6 Gz speaking conditions (speaker 1).

1 G

	ai		ɑʊ	
F1	589	496	533	558
	(0)	(22)	(23)	(44)
F2	1420	1538	1097	930
	(26)	(28)	(56)	(28)
F3	2089	2127	1921	2163
	(28)	(89)	(51)	(172)

+6 Gz

F1	661	548	599	568
	(47)	(18)	(18)	(65)
F2	1168	1478	1033	901
	(78)	(18)	(47)	(93)
F3	2273	2222	2056	2108
	(47)	(141)	(109)	(215)

Table 5. Means and standard deviations, in Hz, of initial and final portion of diphthongs produced in 1 G and +6 Gz speaking conditions (speaker 2).

1 G

	ai		oʊ	
F1	539	463	515	489
	(35)	(26)	(17)	(14)
F2	1414	1705	1085	815
	(130)	(65)	(79)	(41)
F3	2133	2186	2065	2015
	(51)	(48)	(107)	(49)

+6 Gz

F1	612	519	537	548
	(30)	(30)	(47)	(18)
F2	1163	1511	1013	765
	(78)	(53)	(47)	(72)
F3	2209	2255	2046	2211
	(89)	(96)	(107)	(36)

Fundamental frequency. The mean and range of the fundamental frequency, as measured in all syllables receiving primary stress and in unstressed syllables of two-syllable words under the two acceleration conditions, are given in Figures 4a & b. The F0 range of speaker 2, who uses some very low vocal fold vibration rates on the final portions of the last syllables of words, is greater than that of speaker 1. Under acceleration, the mean fundamental frequency value for both speakers is raised and the fundamental frequency range is increased in stressed syllables. For speaker 1, the F0 mean and range in unstressed syllables remain unaffected by acceleration; for speaker 2, the range remains about the same while the mean increases.

Word duration. The mean and range of the duration of each of the five test words are given in Figures 5a & b. The two speakers differ in the durations of words they habitually employ. The effects of acceleration on the duration of words are also different for the two speakers. Speaker 1 speaks relatively rapidly so that his productions of the test words are almost invariably shorter than those of speaker 2. The ranges of his productions are very limited, indicating little variability in word duration. Four of the five test words have a slightly higher mean duration under acceleration; the range of durations does not appear to be affected.

The range of the second speaker's utterances is quite extensive. The effects of acceleration on word duration are variable. For three test words, CCIP, frequency, and two, the mean word duration increases under acceleration; for the

remaining two words, the mean duration decreases. For two words, zero and frequency, the range increases under acceleration; for the remaining words, the ranges decrease.

Segment duration. Of the consonants, only the duration of intervocalic obstruents could be measured with adequate accuracy. The measured consonants were the fricatives, /s/ in the words frequency and CCIP and /v/ in seven, and the stops /p/ in CCIP and /k/ in frequency.

Both speakers showed slight but very systematic differences in the duration of all these consonants; under acceleration, the mean duration of each consonant decreased slightly.

The vowels which could be segmented for unambiguous duration measurements were /u/ in two, /ɛ/ in seven, final /i/ in frequency, and first and final /i/ in CCIP. Acceleration had different effects on vowel durations for the two speakers. For speaker 1, the mean duration of each vowel invariably increased. These increases were as little as 14 msec for the first vowel in CCIP, to as much as 37 msec for the final vowel in the same word. For speaker 2, the effect of acceleration was variable; the average duration of two vowels, /i/ in frequency and /ɛ/ in seven, decreased under acceleration, by 6 msec and 20 msec respectively. The duration of the remaining vowels increased, from 14 msec for the first vowel in CCIP to 90 msec in two.

It would seem that the changes in word duration associated with acceleration are primarily a function of vowel duration, varying from speaker to speaker.

Amplitude. There was no appreciable difference between the 1 G and +6 Gz speaker condition in measures of relative amplitude within the selected high-frequency and low-frequency bands. Examination of LPC spectrum slices for a selection of words produced at 1 G and +6 Gz showed considerable variability but no detectable trends.

SUMMARY AND CONCLUSION

Even though speech produced under high levels of acceleration does not sound unintelligible or distorted, it does exhibit some specific acoustic-phonetic changes when compared with speech produced under 1 G conditions.

The formant pattern was generally affected in the same way for both speakers. The first formant was higher for all vowels except /u/ for speaker 2. The second formant tended to be lower for the front vowels /i, ɛ / and higher for the back vowel /u/. In general,acceleration appears to shift the formants towards a more centralized position, resulting in a more compact vowel space. The formants of diphthongs were also shifted in the same way for both speakers.

Fundamental frequency increased in stressed syllables for both speakers. Even though a concomitant increase in the higher frequency components of the speech spectrum would be expected (i.e. a difference in spectral tilt), it was not observed.

The effects of acceleration on word duration were different for the two speakers. One speaker produced almost all words under acceleration at a greater mean duration; for the other

speaker, some words increased in duration, others decreased. Mean vowel duration followed the same pattern as word duration, increasing for almost all vowels for one speaker, some increasing and others decreasing for the second. Consonants were slightly shorter in words produced under acceleration for both speakers.

These are small but measurable and relatively systematic differences between speech produced under high levels of acceleration and speech produced in a 1 G environment. It is possible to speculate about the changes in speech production which are responsible for these differences.

The consistent formant changes observed under acceleration depend on changes in the articulatory pattern. It seems unlikely that impaired mobility of the articulators is responsible, since fine motor control has been demonstrated to be unimpaired at accelerations higher than +6 Gz. Rather, the changes in articulation may very well be associated with the straining anti-G maneuvers practiced by the speakers. These maneuvers would tend to increase pharyngeal tension and lead to lessened mobility of the articulators, particularly of the tongue.

The changes in word duration may result from changes in the pattern of respiration. Under high acceleration, the rate of exhalation increases. Since speech is produced with a slow, gradual and very controlled exhalation, it is possible that speakers are attempting to control their exhalation rate and in the process are sometimes overcompensating. Respiration patterns, however, provide no explanation for the differential effects of acceleration on consonant vs vowel duration. That

vowels might become longer while a speaker is trying to control exhalation is reasonable. That consonants become slightly shorter needs an articulatory explanation.

The increase in F0 might be associated with increased laryngeal tension, again a result of the straining maneuvers practiced to counteract the effects of acceleration. The increased laryngeal tension should lead not only to increased F0 but to an increase in the higher-frequency components of the glottal spectrum. This increase was not observed.

FURTHER RESEARCH

It is recognized that these results are preliminary in nature and should not be generalized, being based upon only two speakers. However, indications that relatively systematic changes in the acoustic-phonetic structure of speech produced under high levels of acceleration do occur are important. First of all it indicates that problems may well arise if attempts are made to introduce automatic speech recognition technology into the flight environment without taking these changes in speech production into account. For example, the recognizer trained under relatively benign conditions on the ground or in level flight may perform poorly when used under more dynamic flight conditions where the speaker would be under a considerable G-load. Secondly, if the acoustic-phonetic changes that occur are systematic, research defining these changes may lead to the development of compensatory algorithms that will ensure the

successful application of automatic speech recognition technology in the cockpit.

Further research dealing with the effects of acceleration on speech should explore three areas: different acceleration levels, a larger sample of speakers, and a more extensive sample of English segments.

It would be valuable to investigate acceleration levels between 1 G and +6 Gz, to determine whether minimal effects on speech would be detectable at lower accelerations. Lower acceleration levels would also help in separating the effects of acceleration from the effects of the anti-G straining maneuvers.

Since the two speakers responded differently in some cases, the variability of speaker response to acceleration needs to be investigated.

The generalizations which can be made on the basis of the analyzed words is limited. A selection of words which would provide a more detailed inventory of English consonants and particularly vowels would be valuable.

FOOTNOTES

[1]Active duty Air Force personnel who participate in ongoing acceleration investigations. They are volunteers for such investigations and perform normal duty within various Wright-Patterson AFB organizations. All such personnel undergo an intensive medical evaluation prior to their acceptance as panel members. Routine annual physicals are required as long as an individual is a panel member.

[2]These recordings were produced as a part of a speech data base for the development and evaluation of automatic speech recognition systems intended to be used in high performance military aircraft cockpit environments (Anderson and Moore, 1984; Anderson, Moore, and McKinley, 1985). Mr. John Frazier and Mr. Harald Hille of AAMRL were instrumental in the collection of this data base.

[3]SPIRE (Speech and Phonetics Interactive Research Environment) is an audio processing system developed by the Speech Processing Group at MIT. This system provides window-and-menu-oriented graphics interaction, symbolic computation (e.g., program-writing programs) and high-speech numeric processing for synthesis and analysis of acoustic waveforms. Further descriptions of the capabilities of this system can be found in Zue and Cyphers, 1985.

REFERENCES

Anderson, T. R. and Moore, T. J. (1984). "Characterizing requirements of speech I/O for military cockpit environments." Processings of Speech Tech '84, New York: Media Dimensions, 77-80.

Anderson, T. R., Moore, T. J., and McKinley, R. L. (1985). "Issues in the development and use of a speech recognition data base for military cockpit environments." Proceedings of Speech Tech, '85, New York: Media Dimensions, 172-176.

Burton, R. R., Leverett, S. D., and Michaelson, E. D. (1974). "Man at high sustained +Gz acceleration: a review," Aerospace Medicine, 45, 1115-1136.

Davis, J. R., Ratino, D. A., Van Patten, R. E., Repperger, D. W., and Frazier, J. W. (1984). "Performance and physiological effects of multiple, sequential, +Gx acceleration exposures (space plane boost profiles). Aerospace Medical Research Laboratory Technical Report AFAMRL-TR-84-012.

Gillies, J. A. (1965). A Textbook of Aviation Physiology. Oxford: Pergamon Press.

Peterson, G. E. and Lehiste, I. (1960). "Duration of Syllable Nuclei in English." Journal of the Acoustical Society of America, 32, 693-703.

Sharp, G. R. and Ernsting, J. (1978). "The effects of long duration acceleration," In Aviation Medicine, V. I, Physiology and Human Factors, G. Dhenin and J. Ernsting (eds.), London: Tri-Med Books.

Zue, V. W. and Cyphers, D. S. (1985). "The MIT SPIRE system," Proceedings of Speech Tech '85, New York: Media Dimensions, 277-279.

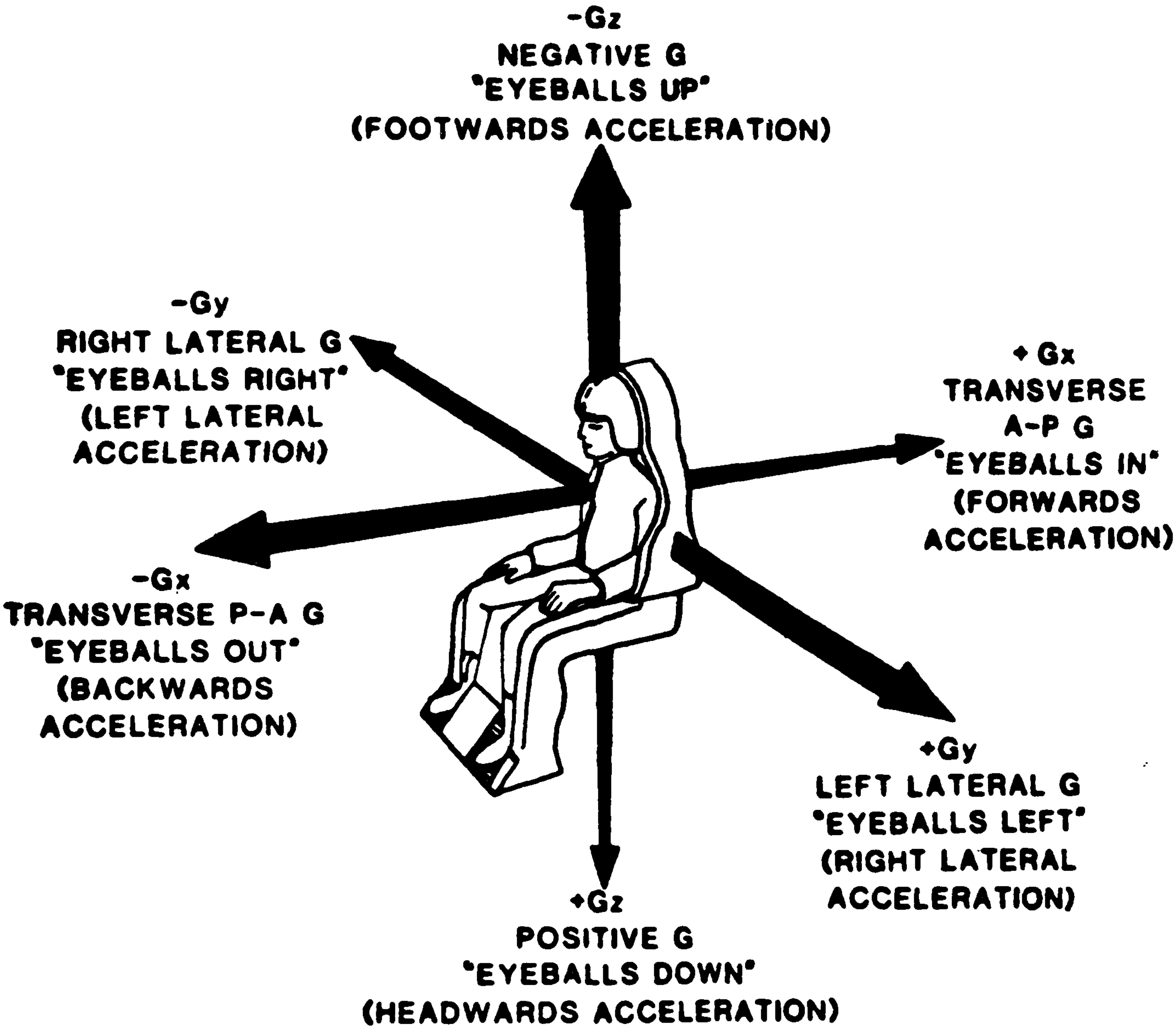


FIGURE 1. G FORCES IN ACCELERATION

Spire Layout 1

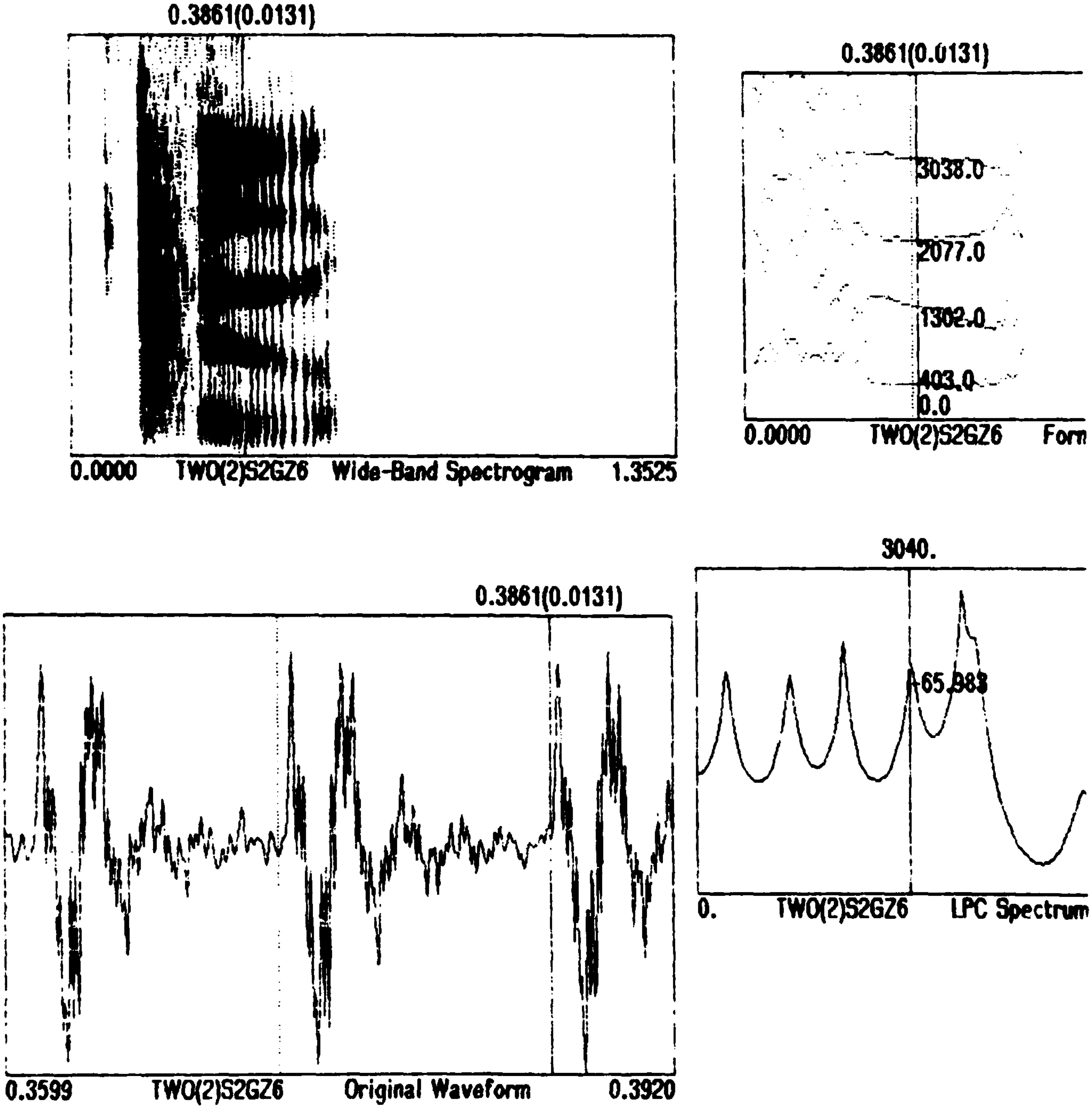


FIGURE 2. DISPLAY PRODUCED BY SPIRE, SHOWING (CLOCKWISE) A WIDE-BAND SPECTROGRAM, FORMANT TRACKS, A WAVEFORM, AND LPC SPECTRUM. ONE GLOTTAL PERIOD IS MARKED ON THE WAVEFORM.

SPEAKER 1

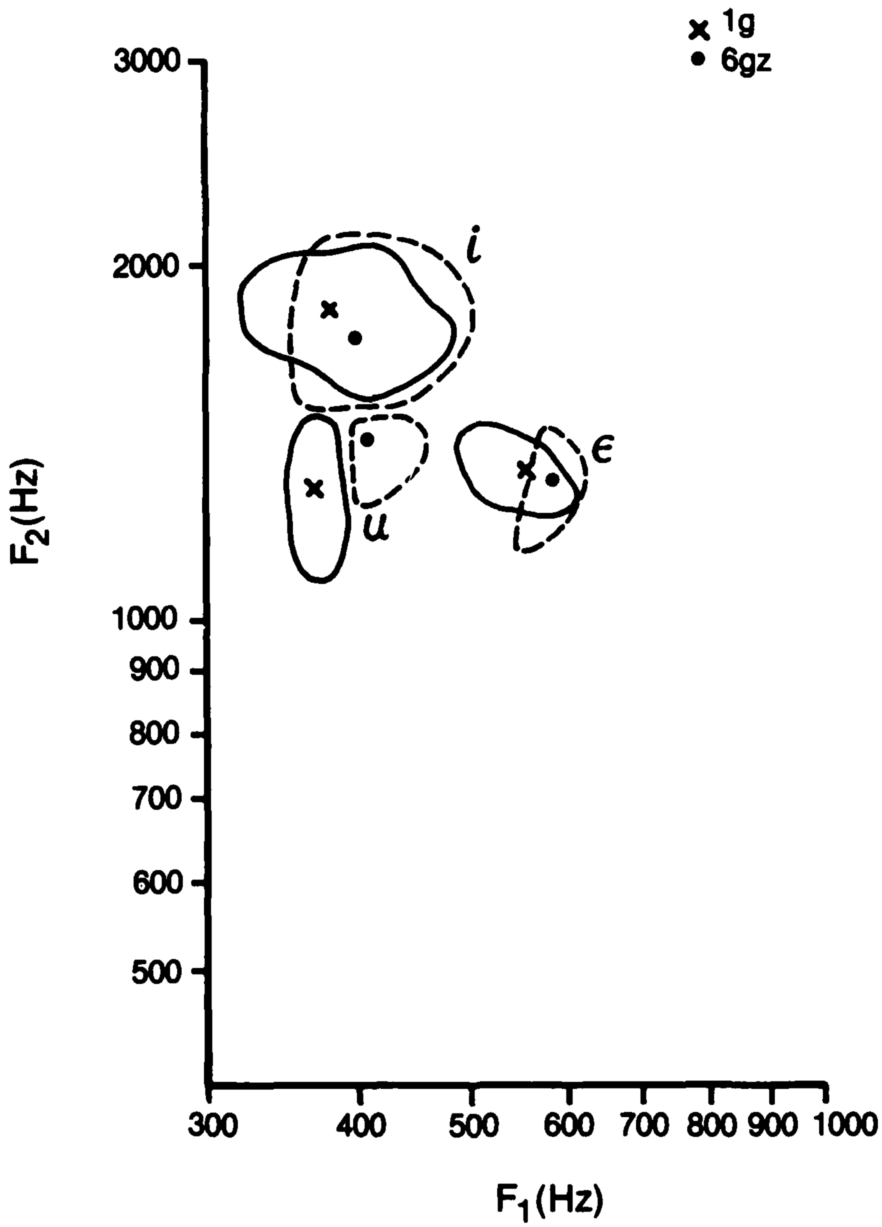


FIGURE 3a&b. The vowel space as defined by F1 and F2, for each speaker. The vowels produced under acceleration at +6 Gz show formant shifts, particularly noticeable for F2. The solid lines form the boundary of the vowel space for vowels produced under 1 G, while the dotted line is the vowel space for the same vowels produced under +6 Gz. The data points represent the mean values for each vowel in their respective vowel spaces.

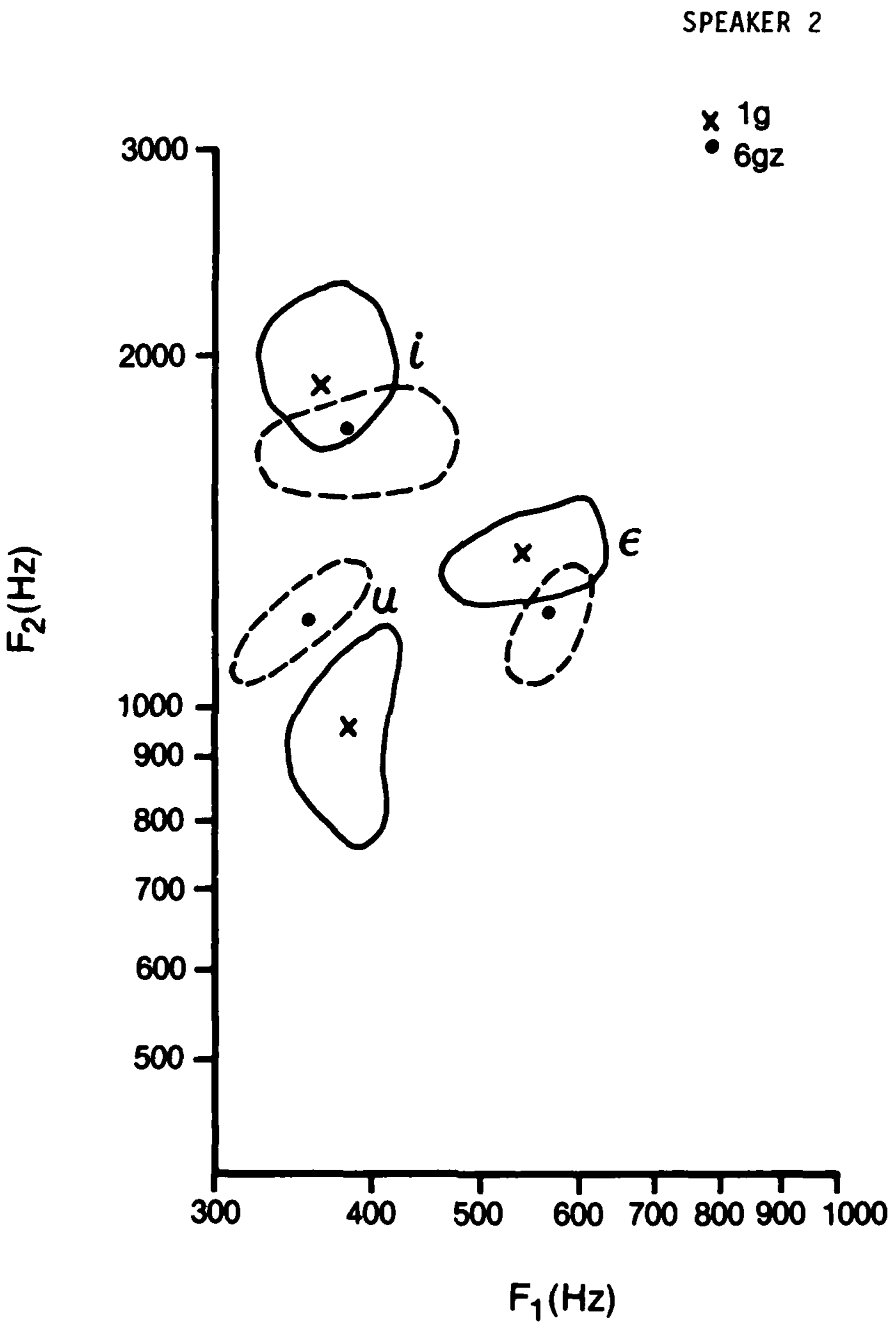

SPEAKER 2
x 1g
• 6gz
3000
2000
1000
900
800
700
600
500
F2(Hz)
i
ϵ
u
300
400
500
600
700
800
900
1000
F1(Hz)

FIGURE 3b.

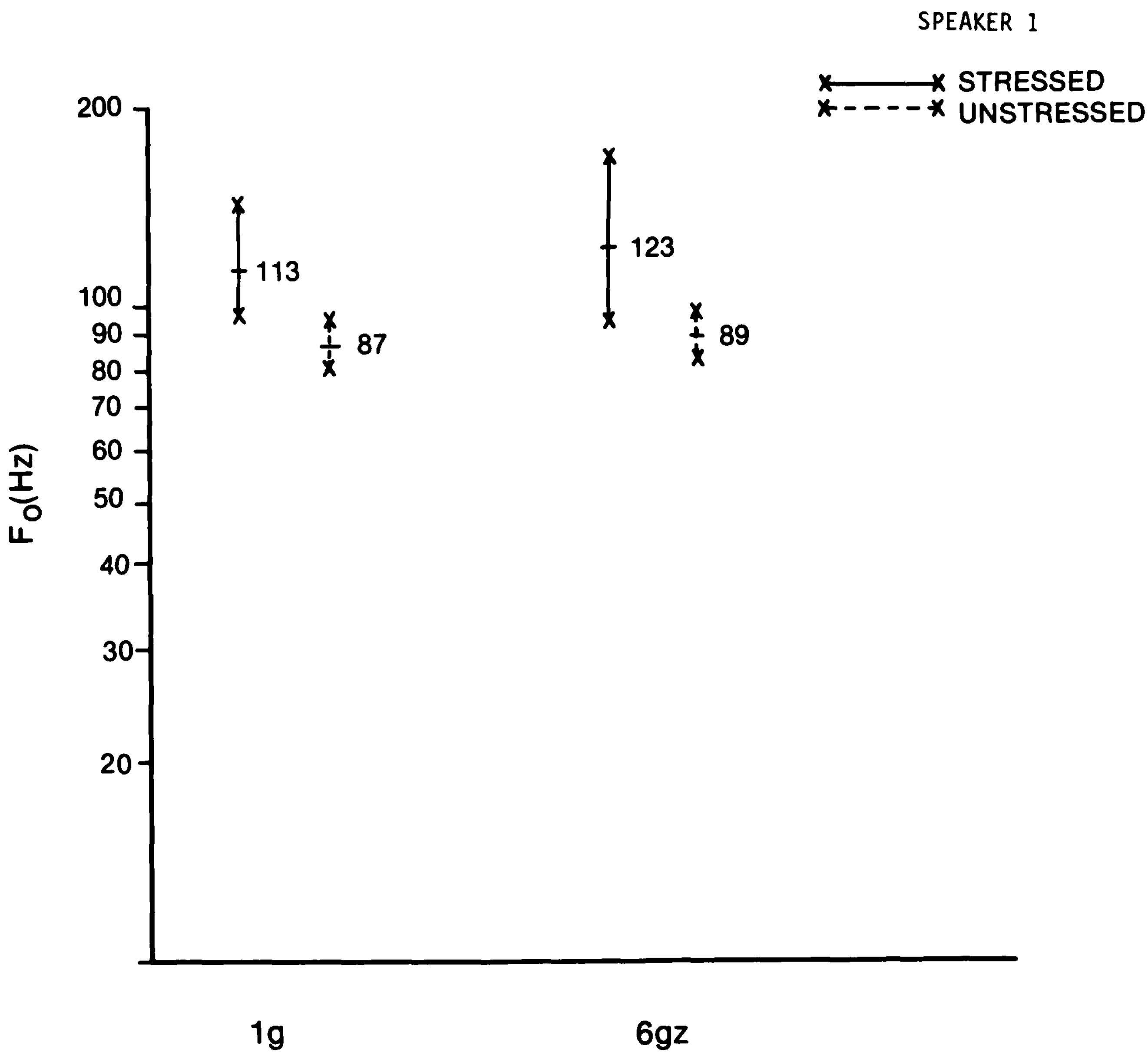


FIGURE 4a&b. Means and ranges of F0 for each speaker at 1 G and +6 Gz.

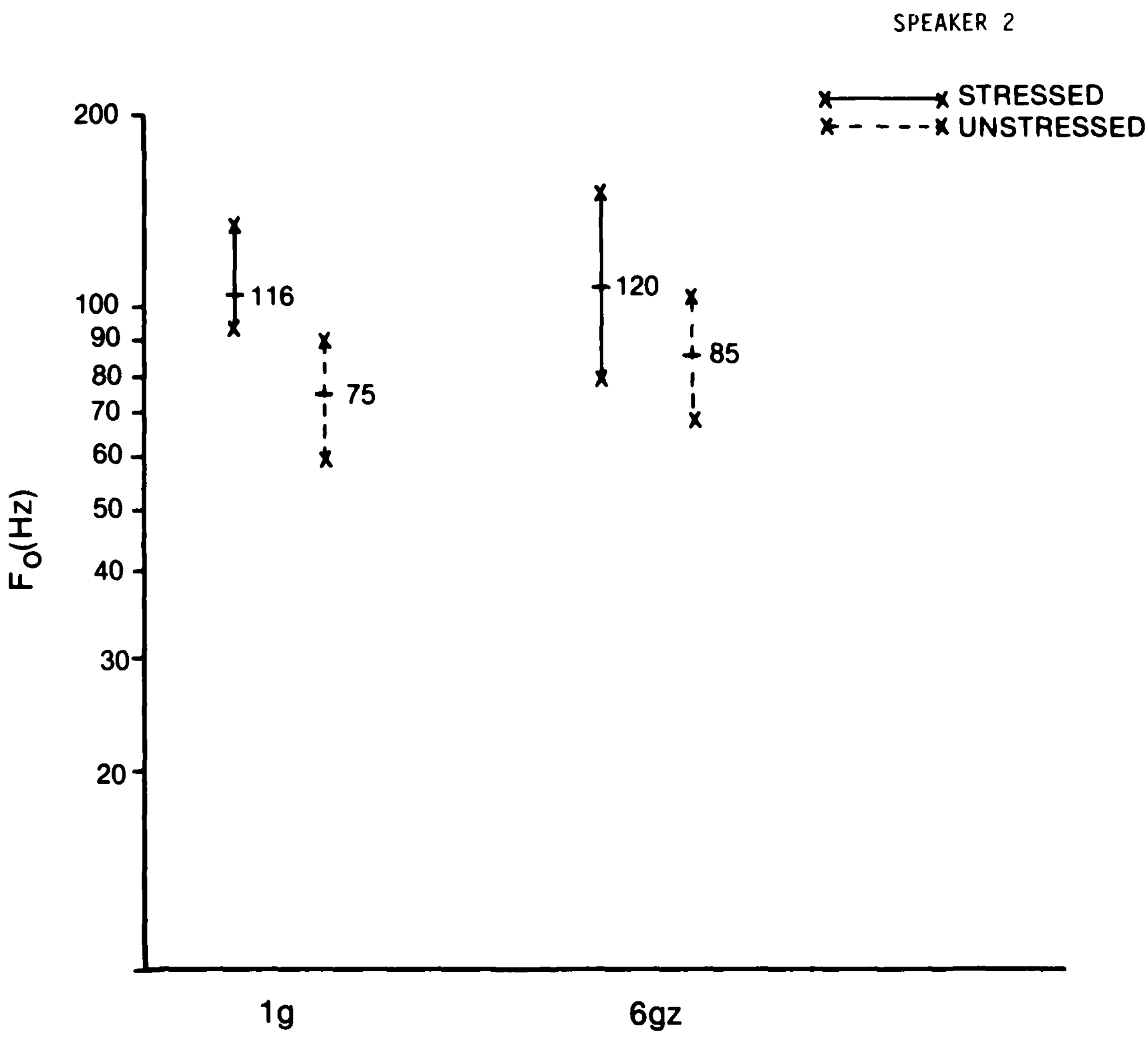

SPEAKER 2
STRESSED
UNSTRESSED
200
100
90
80
70
60
50
40
30
20
F_O(Hz)
116
75
120
85
1g
6gz

FIGURE 4b.

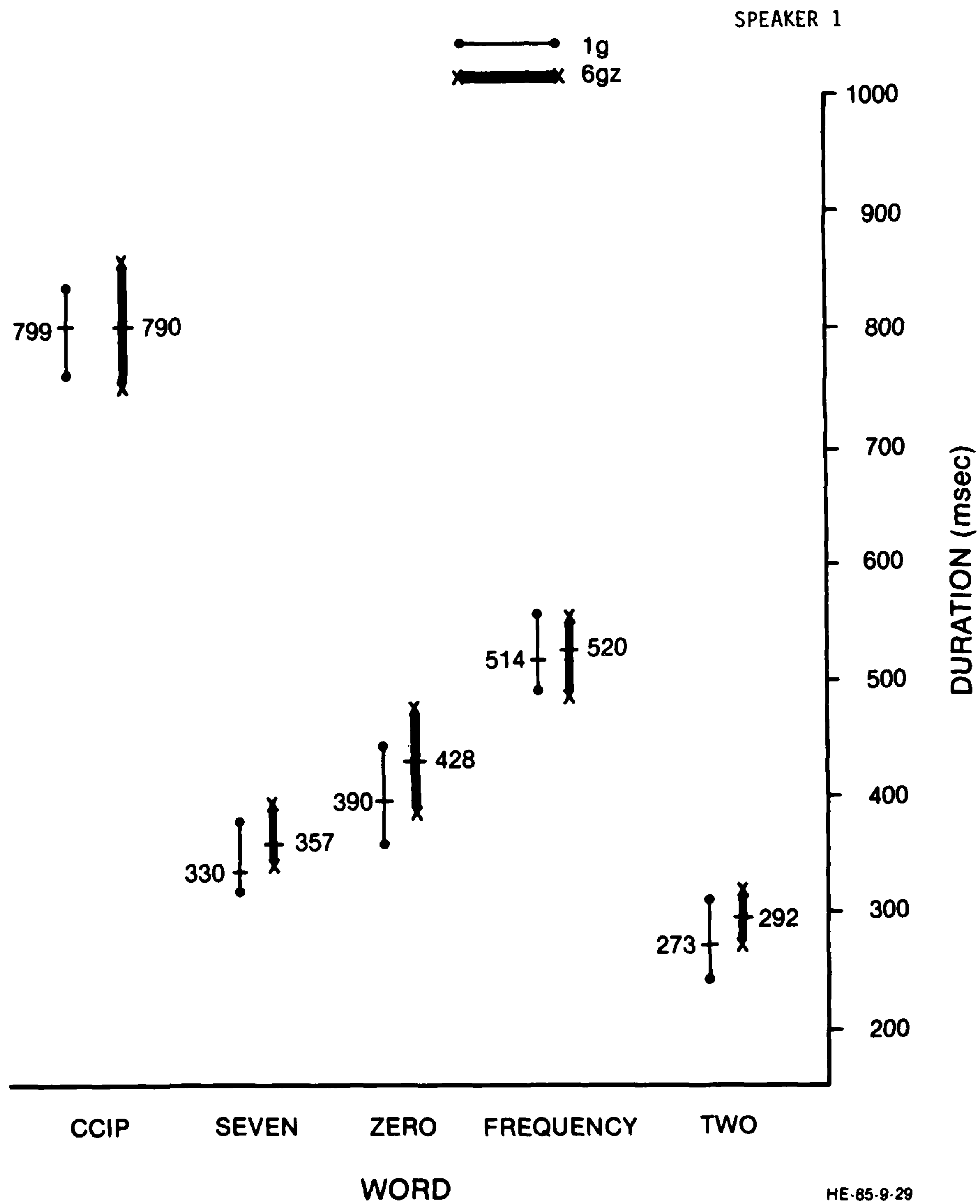


FIGURE 5a&b. Means and ranges of word durations at 1 G and +6 Gz.

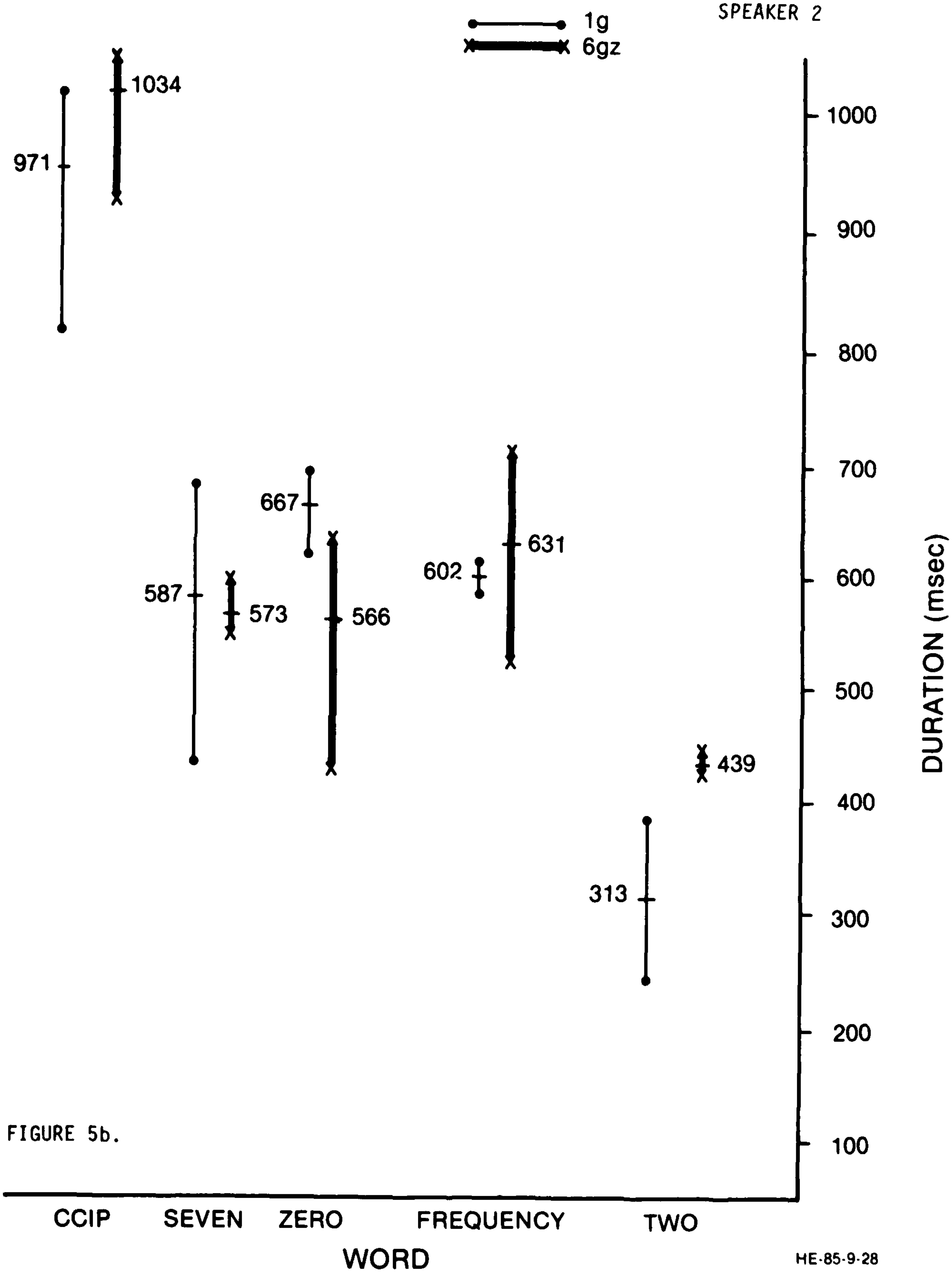


FIGURE 5b.

U S GOVERNMENT PRINTING OFFICE 1988 - 646-006/20236

END
DTIC
9 — 86